



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

UNIFORMITY OF TEACHERS' MARKS VERSUS VARIABILITY

JOHN L. STEWART
High School, Parkersburg, West Virginia

The variability of teachers' marks as recorded in the form of letters or percentages has received a large amount of observation and investigation. Special study has been made of the variation of teachers' marks by having a large number of teachers estimate the same set of papers and evaluate the same set of questions, and by comparing the marks the pupils of various schools receive when brought together in the same school. From these studies it is concluded that the marks assigned by teachers in general are inaccurate measures of the abilities of pupils and are doubtful gauges for determining credit or promotion.

The conclusions of investigators that teachers' marks are an inaccurate means for determining merit must be definitely taken into account by any school which introduces a system of weighted credits. The Parkersburg High School established a policy of awarding six-tenths of a unit for each semester subject in which the pupil received a mark above 92 per cent, five-tenths of a unit for each with a mark between 80 and 92 per cent, and four-tenths of a unit for each with a mark between 71 and 80 per cent. The discussion which follows endeavors to present the effort made by the faculty of this school to discover and to rectify wide variations in marking, in order that the plan of weighted credits might at least approximate justice.

It is reasonable to expect all qualified teachers to give a pupil approximately the same mark for the same recitation unit. Further, it is reasonable to assume that the median ability of a large number of pupils is about the same as the median of any other similar group. Therefore, the median mark of all the pupils of one teacher should be approximately the same as the median

mark of the pupils of any other teacher in the general public high schools under the present classification of pupils.

To test this thesis take the median mark of the reports given out by each teacher; arrange the median marks numerically, and make a graph of the same. Figure 1 illustrates what will be found the first time such a record is made and presented to the teachers. Note Teacher No. 1 had a median mark of 93 per cent; Teacher No. 7, the first quartile, a median mark of 85

Per Cent or
Median Mark



FIG. 1.—Median marks of the teachers of the Parkersburg High School for the first report.

per cent; Teacher No. 21, the third quartile, a median mark of 83 per cent; Teacher No. 27 had a median mark of 75 per cent. Although this graph seems to substantiate the reports of other investigators, the problem will bear further analysis. Certain questions at once present themselves.

Are the pupils of Teacher No. 1 better than those of Teacher No. 27? They were similar groups—therefore the medians should be similar.

Is Teacher No. 1 better than Teacher No. 27? She is considered about the same by the pupils and supervisors.

Are the standards of Teacher No. 1 higher than those of Teacher No. 27? They are not. Teachers seem to be well satisfied with the better pupils in each case.

The wide range in the medians must be on account of the different conceptions of the value of marks. Teachers with like conceptions of values would give similar marks for the same work.

Thus it becomes the principal's problem to unify the teachers' values for the respective marks. This uniformity can generally be brought about by two methods, the one mechanical, the other educational. The mechanical adjustment is a temporary expedient that will bring about a decided uniformity in the marks.

The median as shown by Figure 1 for all the teachers is 84, the quartiles 85 and 83 respectively. According to the thesis, qualified teachers would give the same mark for the same recitation

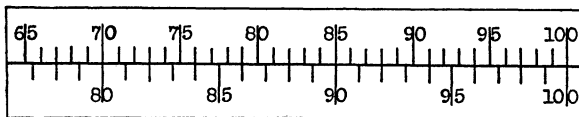


FIG. 2

unit; the median mark of a large number of pupils would be the same as the median mark of any similar group; it is reasonable to request all teachers with a median mark above the first quartile to lower their marks according to a scale; likewise, it is reasonable to have all teachers with a median mark below the third quartile raise all marks according to the scale. Scales are arbitrary; thus they can easily be constructed as in Figure 2. In other words, the pupil receiving the median mark of 75 per cent has a right to expect his mark to be raised to at least the third quartile, or 83 per cent. Thus Teacher No. 27 can readily translate all of his scores into marks that will be in accord with three-fourths of the other teachers. Likewise, Teacher No. 1 can use the scale as a "step down" transformer, lowering his median mark to that corresponding to the median of the teacher of the first quartile, then translating all his scores into marks. This being purely a

mechanical process or arbitrary method, the interesting effect on the teacher's conception of values is to be noted in the comparison of Figure 1 with Figure 3 which gives the second and third reports.

The extremes in medians no longer exist. They have been so modified and the range of variations so narrowed that the pupils now have the satisfaction of electing work or studies for a purpose,

Per Cent or
Median Mark

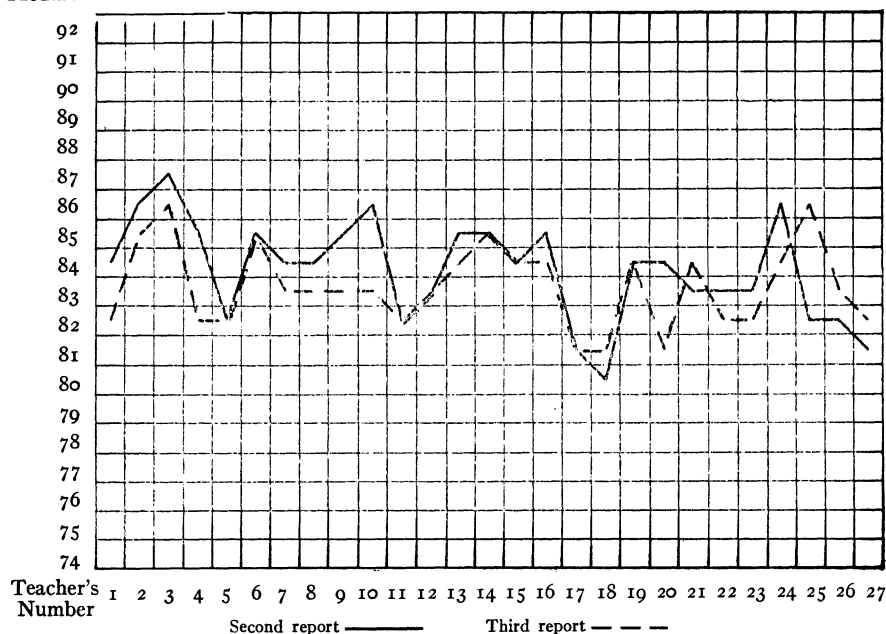


FIG. 3.—Median marks for the second and third reports of the same group of teachers as given in Fig. 1.

rather than the high marks that a particular teacher always gives, or on the other hand avoiding essential work because of a teacher's reputation for giving low marks. The fluctuations of the medians for the second and third reports are but normal and healthy signs that the teachers are not losing their individuality in marking.

The second stipulation, that of education, is fundamental, if any stable or lasting standards are to be established. The proper evaluation of work can probably be best arrived at through a

study of standardized tests and measurements. A study of standardized tests and measurements will not only develop definite requirements for the teacher but will give her an additional opportunity to secure the co-operation of the supervising principal in diagnosing class difficulties. The use of objective standards tends to eliminate the danger of marking or scoring the achievements of a class according to the relative work of the individuals composing that particular group. Further, the teacher acquires a truer conception of satisfactory achievements for each group of properly classified children.

Pupil and class difficulties, or lack of achievements, are diagnosed today by the average teacher or supervisor by methods similar to those of the physician of the nineteenth century in diagnosing a patient's ills. With the incoming of the exact sciences the day of "trial and error" passed in the medical profession. So it is today in the teaching profession with the incoming of scientific methods. Just as our forefathers got well in spite of the quack doctors, they learned the three R's in spite of their schools and teachers. The up-to-date teachers and supervisors of today, like the physician, are scientifically diagnosing the work of the school.

The mechanical or arbitrary method of developing uniformity of teachers' marks is only a temporary expedient that must give way to scientific methods of measuring achievements and diagnosis; the graphic illustration of the median marks of the teachers, however, has a permanent use in helping the new teacher coming into the system immediately to adjust the valuation of her marks to those of the school. The supervising principal, however, cannot be content with knowing only the quartiles and median marks of his teachers; he must know how every teacher arrives at the true valuation of the marks given. Just as our weights, measures, and time are standardized, so must the work in the school be standardized.